DOCUMENT RESUME

ED 117 193 95 TM 005 053

AUTHOR Warren, Jonathan R.
TITLE The Continuing Controversy over Grades. TM Report 51.
INSTITUTION ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO ERIC-TM-51
PUB DATE Nov 75
CONTRACT NIE-C-400-75-0015
NOTE 13p.
AVAILABLE FROM ERIC Clearinghouse on Tests, Measurement, and Evaluation, Educational Testing Service, Princeton, N.J. 08540 (free while supplies last)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Elementary Secondary Education; Evaluation Criteria; *Grades (Scholastic); Higher Education; Reliability; *Student Evaluation; *Validity
IDENTIFIERS Contract Grading; Criterion Referenced Grading; Grade Inflation

ABSTRACT

Although grades have been criticized for lack of reliability, end-of course grades and grade-point averages are reliable enough for most uses. The charge of unreliability applies only to grades on themes, tests, or other individual pieces of student work. On the other side of the controversy, grades have been said to be essential to the learning process because they provide for the evaluation of student performance. But performance is evaluated and its results reported to students independently of any grading system. The justification for grades must lie elsewhere. The critical issue in grading is the validity and usefulness of grades for the variety of purposes they are called on to serve--conveying information on student achievement, providing incentives for students to study, serving as selection criteria, providing material for administrative records, helping in the evaluation and monitoring of the instructional process, and assisting students in educational and occupational planning. Until better information is available on the effectiveness of grades with respect to these various functions, the continued trading of unsupported assertions about them will be fruitless. New approaches to grading, such as contract and criterion-referenced grading, do not change the basic issues. (Author/RC)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

EDU/193

ERIC CLEARINGHOUSE ON TESTS, MEASUREMENT, & EVALUATION
EDUCATIONAL TESTING SERVICE, PRINCETON, NEW JERSEY 08540

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

TM REPORT 51 NOVEMBER 1975

THE CONTINUING CONTROVERSY OVER GRADES

Jonathan R. Warren

Abstract

Grading procedures in schools and colleges have been a topic of controversy since soon after the turn of the century, but little has been learned during that period to resolve the issues in dispute. Although grades have been criticized for lack of reliability, end-of-course grades and grade-point averages are reliable enough for most uses. The charge of unreliability applies only to grades on themes, tests, or other individual pieces of student work. On the other side of the controversy, grades have been said to be essential to the learning process because they provide for the evaluation of student performance. But performance is evaluated and its results reported to students independently of any grading system. The justification for grades must lie elsewhere. The critical issue in grading is the validity and usefulness of grades for the variety of purposes they are called on to serve - conveying information on student achievement, providing incentives for students to study, serving as selection criteria, providing material for administrative records, helping in the evaluation and monitoring of the instructional process, and assisting students in educational and occupational planning. Until better information is available on the effectiveness of grades with respect to these various functions, the continued trading of unsupported assertions about them will be fruitless. New approaches to grading, such as contract and criterion-referenced grading, do not change the basic issues.

TM005 053

The material in this publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to qualified professionals for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either these reviewers or the National Institute of Education.

The questions raised about grading practices are apt to include the following:

What should grades represent?

Should student effort be reflected in grades?

Are grades adequately reliable?

How can their reliability and fairness be improved?

What set of symbols should be used?

Questions similar to these can be found in many educational books and journals of the 1970s. Those listed above were taken from a monograph published more than 60 years ago (9). Answers to the questions and disagreement about the answers have also changed little during the intervening years. Finkelstein (9) and Ebel (8) agree that grades should be based only on student accomplishment. On the other hand, grades are used to reflect pupil growth, aptitude, effort, and attitudes as well as achievement (25). They should reflect achievement in relation to ability, according to Kindsvatter (13); they should not according to Terwilliger (23), Ebel (8), and others. Still others would permit grades to represent both effort and achievement and perhaps other qualities, but independently of one another, (for example, 10), while Ebel (8) points out problems that the use of multiple criteria adds to the already complicated process of providing useful grades.

More than 60 years ago, Kelly (12) reviewed studies available at that time of the variability of grade distributions across colleges and universities, across fields within institutions and across faculty members within fields and concluded that "a given grade or mark means many widely different things to different teachers." The same concern is an issue today (21). Clear specification of the grading criteria - whatever they may be - and the meaning of the points on the particular grade scale used are stated as major requirements for the improvement of grading (3, 7, 8, and 10).

The continued vehemence of the controversy over the grading process without apparent progress after more than half a century is puzzling. Although grading is in some respects a complex, shifting issue, difficult to come to grips with, about 65 years of study and debate would ordinarily be expected to clarify or redefine the issue at least, if not produce some clear advances in understanding and practice. Hiner (11) accounts for the lack of progress by ascribing to the grading process the characteristics of a cultural ritual that serves to reduce the impact of a fundamental social problem. The problem eased by the grading ritual, according to Hiner, is the distribution of rewards in a society that simultaneously values achievement and equality. Grades, in acting as a buffer between two

conflicting values, will necessarily be a source of conflict until the competing cultural values are changed. Some such culturally based interpretation as Hiner's seems necessary in view of the persistence and ardor of the grading controversy and the frequent absence from the arguments of anything more than assertions of belief.

Those in favor of drastically modifying or even abolishing current practices claim that grading interferes with learning by creating anxiety in students as well as a distaste for school, by forcing them to spend their time and energy on grades rather than learning, and by making teachers and students antagonists instead of allies in the learning process. Those who defend grading say it is essential to learning because students would not work as hard in the absence of grades and because effective learning requires the knowledge of results that grades provide. The pro-grade faction argues further that grades are necessary in the selection and placement of students in successively higher levels of education. Both positions are based on the asserted effects of grades on learning, but the common ground on which they can be compared is limited. While the two factions agree, for example, that grading increases the level of competition in learning, the pro-grade faction approves of competition and the anti-grade faction disapproves of it. Neither faction has pursued the implications of competition far enough to be forced to resolve their apparently comfortable conflict.

Five years ago, a review of the growing literature on grading (24) showed the situation to be much the same as it is today. Despite a continuing flood of reports on experimental grading procedures and frequent broadsides from both camps, only one advance in understanding has been made in the five years since the earlier review. The evidence is now overwhelming that pass-fail grading does not induce students to take courses they would have avoided under a traditional grading system for fear of depressing their grade-point average. Proponents of pass-fail grading have therefore lost one of their major positions, but one that seems not to be critical in the broader controversy.

Despite continuing moderate interest in pass-fail grading, on the college level it is rarely allowed to apply to any more than a minimal part of a student's program. It is less common still in high school (19). In elementary school, where it more commonly appears as S-U (satisfactory - unsatisfactory), its impact is modified by the use of several rating scales, which gives the appearance of providing enough information to satisfy the users of grades. The major impact of pass-fail grading in college seems to be its contribution, to an unknown degree, to the inflation of undergraduate college grades.

Grade Inflation

Since the late 1960s, college grade-point averages have risen by about half a letter grade (20); the most common college grade is now a B rather than a C. Possible contributing factors other than an increase in pass-fail grading, which may have reduced the number of Cs and Ds proportionately more than the As and Bs, include the refusal of faculty members to give low grades to college men when satisfactory grades could keep them from being drafted; a growing dissatisfaction with the grading process as a whole; increasing frequency of field experience, internships, and other nontraditional kinds of courses in which grades below a B seem to be rare; and, to a small extent, the recently growing practice of nonpunitive grading--the abandonment of Fs in favor of simply not reporting a grade. While each of these may have contributed to the rise in grade-point averages over the past decade, the proportion of As and Bs relative to all other grades, including pass, has grown substantially. Pass-fail and nonpunitive grading cannot have contributed to that increase.

The growth of nonpunitive grading (either pass-no report or ABCX, with X indicating that no grade was being reported) followed pass-fail grading by a few years. Arguments in its favor seemed overwhelming in comparison with arguments against it. The primary opposition was based on one fear that permitting students to take a course and fail it, without keeping a record of their failure, would lead to an overburdening of college resources by poorly prepared or frivolous students who would crowd classes in search of a few in which they might succeed, denying places to better prepared or more serious students. Yet there are simple procedures, such as requiring students to maintain a minimal rate of course completion for continued enrollment, that would prevent abuse of a no-fail grading system. Other reasons for opposition have been the desire to prevent the range of the grade-point average from shrinking by maintaining the bottom grade and the belief of some faculty members that failure is a phenomenon from which students should not be sheltered. None of these objections to the removal of failure from the grading process is compelling, yet recent shifts to nonpunitive grading are being reversed.

While college grades have clearly risen since the mid or late 1960s, the recent furor over grade inflation seems to be reversing that trend - a trend that never was universal (14). The desire of graduate and professional schools to restore greater spread to the undergraduate grade-point average for greater ease of selection and the desire of faculty members and some students to maintain a system of public recognition of superior performance are probably the primary reasons behind current efforts to roll back the recent rise in college grades.

Purposes of Grades

Basic to any consideration of the good or evil that results from the use of grades is an understanding of their purposes and of the ways the grading process accomplishes them. There is general agreement on the purposes themselves, but misunderstanding has persisted over how they are achieved.

At precollege levels of education, one of the primary purposes of grades is to report to parents on the school performance of their children. Many writers consider this the most important function of grades at precollege levels, but at the college level that purpose almost disappears. At the elementary level, the most common form of grade is the dichotomous Satisfactory-Unsatisfactory applied to a varied group of qualities that may include knowledge, growth, potential achievement, effort, attitudes, conduct, attendance, citizenship, or character (8, 25, and 26). The use of checklists showing each of the qualities to be assessed and reported is fairly common in elementary schools but less so at the secondary level (1 and 19). In college it is quite rare. With each advancing educational level, then, the grading process shifts from multiple criteria, each graded S or U, to a single criterion - academic achievement - graded at several levels. Other considerations, such as effort, attitude, or interest, enter the process to a degree known only to the individual teachers assigning the grades. With this shift in the form of grading used at successively higher educational levels, the amount of information conveyed by grades in a formal sense is reduced, but the amount needed has dropped as well. At the college level grades are rarely used to provide information to parents, and the students already have most of the information in grades. The student who does not know in advance with reasonable accuracy what his grades will be is rare.

The second commonly cited purpose of grades is to induce students to study. The nature of the inducement and its effect, however, are complicated. Achievement orientation, age, and the degree of competitiveness in the testing situation can be expected to interact in their effects on student performance, reports McKeachie (15). Other studies he reviewed showed that the effects of teachers' evaluative communications to students depended on whether they were primarily informational or contained an element of praise.

A recent study of high school students showed the threat of low grades to be a greater inducement to study than the appeal of high grades, and both to be more effective than no grades at all (5). These results raise new questions about the move in some colleges and universities to abolish failing grades. But the authors warn that a large number of complicating factors probably affect the role of grades as incentives.

At the elementary level grades probably serve pupils almost

entirely as indicators of adult approval and therefore of personal worth. As the educational level advances, grades take on additional meaning as indicators to students that they have accomplished something of value that is independent of adult approval. The importance of grades to parents also increases with each grade level, which affects their importance to students in complicated ways. Furthermore, the higher the educational level, the more important grades become as a valued commodity in their own right, as some high school students become concerned about admission to college and college students begin planning for graduate or professional school. The motivating function of grades therefore varies, depending on the student's educational level, his or her intentions and inclinations, family relationships, personality, and other characteristics of student and school.

At the college level, and to some extent at the secondary level, grades are used for selection and placement. High school grades determine selection by particular colleges and universities and college grades determine selection by graduate and professional schools. In high school and junior high school as well as college, grades help determine whether students are placed in select, regular, or remedial classes. As students progress, this function of grades becomes more critical. In secondary schools, placement may rest primarily on the judgment of individual teachers, with grades acting only as prompters. Faculty judgment operates somewhat less strongly in placement in college classes but still has some effect in admission to college and to graduate and professional schools through faculty recommendations. In these selection and placement functions, grades operate as summary statements of a number of teacher judgments.

The function of grades as concise indicators of teacher judgments gives them an important role as the primary criterion against which selection and placement decisions are validated. Both types of decision are intended to assure that students will be engaged in programs at a level at which they can succeed. Whether based on prior grades, teacher recommendations, test scores, prior school experience, or other considerations, the success of selection and placement decisions is usually determined through reference to grades achieved. Other criteria could be used and for many purposes would be preferable to grades, as when the purpose of selection is to assure the presence in a program of students who will benefit most, in some specified sense, from the particular program offered. But grades are familiar, convenient, and almost universally accepted as the dominant criterion for successful selection and placement. Standardized tests of academic aptitude and achievement depend almost exclusively for their acceptance on the degree to which they are related to grades.

Providing information to parents, students, and others about the level of student achievement, providing a basis for admission to more advanced educational programs, and serving as inducements to

study are the three most commonly agreed on purposes of grades. Others include helping students choose educational and occupational goals, monitoring the effectiveness of instructional programs, and providing information for administrative records for purposes of promotion, awards, probation, dismissal, and other administrative decisions.

Points of Confusion in Grading

The reliability of grades, which Finklestein and others (9 and 23) questioned more than 60 years ago, is still a source of confusion because of uncertainty about which point in the grading process is at issue. At an early point in the process, individual instances of student performance--tests, quizzes, written reports, essays or themes--are evaluated and graded. Later, the teacher combines these grades with impressions he has derived from informal observations of the student during the course of a school term to arrive at an overall grade for the course. While there is still disagreement as to whether qualities like students' attitudes, interest in the course, or effort should enter into course grades, the weight of opinion favors limiting course grades to reporting academic achievement. Nevertheless, faculty members, even in advanced college courses, are often reluctant to give the same grade to two students equal in achievement, one a very bright student who coasted without effort and with little interest through the course and the other a less bright but conscientious student who worked hard for what he or she learned in the course. Finally, course grades are combined into grade-point averages for use in selection by colleges and universities, graduate and professional schools, honorary societies and scholarship and fellowship awards committees. Despite substantial differences in the characteristics of grading for each of these, the distinctionss are often ignored in discussions of grades.

One of the important differences among grades on particular student products, course grades, and grade-point averages is their reliability, an issue raised by Finkelstein (9). The judgments made by teachers in assigning grades to individual pieces of work are frequently unreliable in the sense that the same teacher may not assign the same grade to the same piece of work if judged at different times, and different teachers often would assign different grades to the same product. This is usually the part of the grading process people refer to when they cite the unreliability of grades. But overall course grades are reliable or consistent enough across faculty members and in the course of reasonable periods of time to be useful measures of achievement. And grade-point averages are more reliable still (2 and 4). They are generally reliable enough to justify their use in admission decisions about individual students; and many of the alternative selection criteria, such as recommendations or judgments based on interviews, are far less reliable. Course grades are also reasonably reliable indicators of comparative levels of student achievement. Even the grades for

individual pieces of work can be made reliable through clear definition of the qualities to be judged and careful observation and assessment of the evidence on which the judgments are to be based.

Validity

The issue on which grading, as it is typically carried out, is vulnerable to criticism is not reliability but validity. As predictors of future grades, grades are about as valid as anything else. But the prediction of future grades is a limited basis for deciding on the validity or usefulness of grades. The variety of purposes grades are intended to serve implies numerous other validities for grades. A procedure with relatively low reliability that is nonetheless valid for a particular purpose is preferable to a more reliable measure that is unrelated to the behavior of interest. The same process is unlikely to serve all purposes equally well.

The different uses and validities of grades are the aspect of grading most in need of study and most subject to unsupported assertions that grades are either good or bad. Studies of grading processes have provided evidence on the validity of grades for particular purposes, usually admission, but they are too narrow in their focus to provide much knowledge about grading as a major educational institution, or, in Hiner's (11) view, as a cultural ritual. The value of grades in all their functions should be studied if the controversy is to continue.

Evaluation and Grading

Another source of confusion in considerations of grading is the distinction between evaluating a student's performance and reporting the result of that evaluation in the form of a grade. A familiar evaluation process is the one in which a teacher reads a student's paper and judges her perception in bringing certain arguments to bear, her grasp of the underlying issues, the coherence of her presentation, the soundness of the conclusions she reaches, and the elegance and correctness of her prose. The teacher can evaluate each of these aspects of the paper and write extensive comments to the student without any concern for grading.

Because grades are typically required at the end of a course, after the evaluation process just described, the teacher will probably estimate how the paper compares with those of other students in comparable courses and assign a grade based on that comparison. But evaluating the student's work and discussing it with her - the parts of the grading process that are so important to learning - can take place without any concern for grading at all. This confusion between evaluation and grading is the source of many vehement assertions in the literature that grading is essential to student learning. Evaluation is essential; grading is not.

Maintaining the distinction between evaluation and grading also clarifies discussions of the effect of grades on motivation. Grading itself clearly has consequences for students' motivations, as when the fear of receiving less than an A induces a premedical student to study for a biology test instead of immersing himself in Dostoyevsky's Notes from Underground. But the motivational consequences for students of knowing that their work will be critically evaluated are as clearly of a different sort and have little if anything to do with grading.

Criterion-referenced and Contract Grading

Although criterion-referenced and contract grading are not the same, they are closely related. Furthermore, they are both relatively new considerations in discussions of grading.

In criterion-referenced grading, the performance expected - both kind and level - is specified in a way that permits its accomplishment by each pupil to be observed independently of the performance of others (17). Thus "Adds three single-digit numbers" is a kind of performance that can be observed in an individual pupil without reference to other pupils. That performance plus others in arithmetic and in other subjects can be listed and checked off as students learn to do them, providing a comprehensive record of each pupil's achievement. Used as a report card, such a checklist would be similar to those advocated by Wrinkle (26) and in use for about 10 percent of elementary pupils (1). The element that is new and distinctive is the specification of the criterion of performance accomplished instead of the judgment that performance is satisfactory.

Contract grading occurs when the student and faculty member, usually in college but occasionally in high school, agree at the beginning of a course on the amount and quality of work the student must do to earn a given grade (6, 16, and 18). The contract may be written so that its completion can be determined unambiguously - for example, set up and carry out an analytic procedure using specified processes to identify the chemical impurities and their concentrations in a sample of river water. When a student has met such specified performance criteria, he or she has completed the course and will receive the agreed-upon grade. Contract and criterion-referenced grading can thus be readily combined. Contract grading does not necessarily require the statement of precise criteria, but problems are avoided if it does. In any case, the procedures for evaluating the student's work and the quality required, if the contract is to be satisfied, should both be specified.

When used sensibly, both criterion-referenced and contract grading can be effective devices for indicating student achievement but they are not always easy to use. Even at the elementary level, where academic requirements can often be clearly and simply stated,

as in "Adds three single-digit numbers," problems in assessment and reporting are not totally avoided. How many times must second-graders, for example, add three single-digit numbers without error to be considered proficient? How long must they maintain their proficiency? Some ninth-graders have learned to multiply simple fractions successfully every term since the sixth grade, yet start a new class in arithmetic having to learn it once more. A teacher's judgment of proficiency may reflect subtle yet important aspects of performance that specifically stated criteria may miss. While the criteria can be worded to accommodate questions of persistence in learning and other complexities, they cannot reasonably be elaborated indefinitely.

The value of both criterion-referenced and contract grading is in their requirement that educational objectives be carefully examined and clearly defined, and that procedures for assessment be carefully worked out in advance. In contract grading, the participation of the student in establishing the evaluation procedure as well as the course objectives minimizes the common student complaint that the tests in a course were unfair.

The Present Status of the Grading Controversy

Each of the questions raised long ago by Finkelstein is still alive, although whether grades are reliable or not no longer deserves to be. Many grades on individual pieces of student work are quite unreliable; course grades are usually reliable enough for many purposes; grade-point averages are quite reliable. What grades should represent is a more complex issue than whether they should reflect effort or attitude. Disagreement still exists on effort and attitude, although consensus is against them. Consideration of the kinds of performance grades should represent cannot be separated from consideration of the purposes for which the grades will be used, an issue that complicates questions of validity and bears on Finkelstein's concern for fairness. What set of symbols should be used may be a more lively issue today than in Finkelstein's time. It is usually concerned today, as then, with the number of points to be included on the grading scale, although instances still occur in which nothing is changed but the label - High Honors, Honors, Low Honors, Pass, and Fail as a replacement for ABCDF.

Studies of grades and surveys of grading practices are numerous and pointlessly repetitious. The most promising way to get off the grading merry-go-round is to focus studies of grades on their several purposes and on alternative ways of accomplishing them. If the questions to be pursued are selected and phrased to be critical to both sides of the controversy rather than simply confirming a point that is not in serious dispute, some advance may be made. But if grades do indeed serve unacknowledged functions, such as easing social conflict or giving insecure or punitive faculty members a device for controlling students, even sound, well-executed studies will not resolve the controversy.

REFERENCES

1. Austin, D.B. Grading systems. In L. C. Deighton (Ed.), The encyclopedia of education. Vol. 4. New York: Macmillan and Free Press, 1971. Pp. 181-185.

2. Barritt, L. S. Note: The consistency of first semester college grade point average. Journal of Educational Measurement, 1966, 3, 261-262.

3. Brown, F. G. Measurement and evaluation. Itasca, Ill.: F. E. Peacock, 1971.

4. Clark, E. L. Reliability of grade-point averages. The Journal of Educational Research, 1964, 57, 428-430.

5. Cullen, F. T., Jr., Cullen, J. B., Hayhow, V. L., & Plouffe, J. T. The effects of the use of grades as an incentive. The Journal of Educational Research, 1975, 68, 277-279.

6. Dash, E. Contract for grades. Washington, D.C.: ERIC Clearinghouse on Higher Education, 1970.

7. Davis, F. B. Educational measurements and their interpretation. Belmont, Calif.: Wadsworth, 1964.

8. Ebel, R. L. Essentials of educational measurement. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

9. Finkelstein, I. E. The marking system in theory and practice. Educational Psychology Monographs No. 10. Baltimore: Warwick & York, 1913.

10. Gronlund, N. E. Improving marking and reporting in classroom instruction. New York: Macmillan, 1974.

11. Hiner, N. R. An American ritual: Grading as a cultural function. The Clearing House, 1973, 47, 356-361.

12. Kelly, F. J. Teachers' marks: their variability and standardization. New York: Teachers College, Columbia University, 1914.

13. Kindsvatter, R. Guidelines for better grading. The Clearing House, 1969, 43, 331-334.

14. Kramer, R., Kaufman, B., & Podell, L. Distribution of grades: 1972. New York: Office of Program and Policy Research, City University of New York, 1974.

15. McKeachie, W. J. Instructional psychology. In M. R. Rosenzweig & L. W. Porter (Eds.) Annual Review of Psychology. Palo Alto, Calif.: Annual Reviews, Inc., 1974, 161-193.

16. Miller, G. D. Grades for reinforcing independent thinking. Improving College and University Teaching, 1974, 22, 175-181.

17. Millman, J. Reporting student progress: A case for a criterion-referenced marking system. Phi Delta Kappan, 1970, 52, 226-230.

18. O'Kane, J. M. Having students do the grading. Improving College and University Teaching, 1971, 19, 331-332.

19. Pinchak, B. M. & Breland, H. M. Grading practices in American high schools. The Education Digest, 1974, 39(7), 21-23.

20. Scully, M. G. Grade inflation. The Chronicle of Higher Education, Oct. 7, 1974, 9(3), 2.

21. Soderberg, L. O. Grading: An old problem revisited. Improving College and University Teaching, 1974, 22, 244-246.

22. Starch, D. Reliability and distribution of grades. Science, 1913, 38, 630-636.

23. Terwilliger, J. S. Assigning grades to students. Glenview, Ill.: Scott, Foresman, 1971.

24. Warren, J. R. College grading practices: An overview. Report 9. Washington, D.C.: ERIC Clearinghouse on Higher Education, 1971. ED 047 164.

25. Williams, R. G., & Miller, H. G. Grading students: A failure to communicate. The Clearing House, 1973, 47, 332-337.

26. Wrinkle, W. L. Improving marking and reporting practices in elementary and secondary schools. New York: Rinehart, 1947.

Items followed by an ED number (for example, ED 047 164) are available from the ERIC Document Reproduction Service (EDRS). Consult the most recent issue of Resources in Education for the address and ordering information.

70774 • V115P2 • 275317 • Printed in U.S.A.